AD-A068 145 TEXAS A AND M UNIV COLLEGE STATION INST OF STATISTICS F/G 12/1

EMPIRICAL STUDY OF QUANTILE REGRESSION ESTIMATORS.(U)

FEB 79 M MELZER DAAG29-78-G-0180

UNCLASSIFIED TR-A-2 ARO-16228.3-M NL

1 OF 1

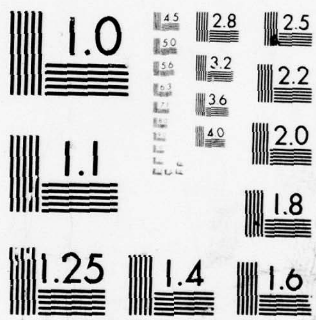
AD A068145

END

DATE FILMED

6 --79

DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD A068145

INSTITUTE OF STATISTICS
Phone 713 - 845-3141

TEXAS A&M UNIVERSITY
COLLEGE STATION, TEXAS 77843

EMPIRICAL STUDY OF QUANTILE REGRESSION ESTIMATORS

By Michael Melzer

Technical Report No. A-2

February 1979

Texas A & M Research Foundation
Project No. 3861

"Maximum Robust Likelihood Estimation and Non-parametric Statistical Data Modeling"

Sponsored by the U.S. Army Research Office

Professor Emanuel Parzen, Principal Investigator

Approved for public release; distribution unlimited.

79 04 20 031

(18) ARO (19) 16228.3-M

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE — READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Technical Report No. A-2	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): (6) Empirical Study of Quantile Regression Estimators		5. TYPE OF REPORT & PERIOD COVERED: (9) Technical rept.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): (10) Michael Melzer		8. CONTRACT OR GRANT NUMBER(s): (15) DAAG29-78-G-0180
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Texas A&M University, Institute of Statistics, College Station, TX 77843		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Army Research Office, Research Triangle Park, NC 27709		12. REPORT DATE: (11) February 1979
		13. NUMBER OF PAGES: 54
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): (12) 29p.		15. SECURITY CLASS. (of this report): Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

(14) TR-A-2

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

NA

18. SUPPLEMENTARY NOTES

The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Nonparametric regression, quantile regression, kernel estimators, statistical data modelling

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Given bivariate data $((X_i, Y_i)\ i = 1, \ldots, n)$ one would like a technique for estimating the regression function $r(x) = E(Y|X)$, which is applicable for a wide range of functions, $r(x)$. Further, one would like a technique with good precision over the full range of values for X.

Techniques of quantile regression based on the approach developed by Parzen (1977), and discussed by Carmichael (1978), will be compared to each

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

347 380

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

other and to Benedetti (1975) in the case where X is restricted to the unit interval. Further, the quantile regression techniques will be compared in the case where X is a random variable.

ACCESSION for		
NTIS	White Section	☑
DDC	Buff Section	☐
UNANNOUNCED		☐
JUSTIFICATION		
BY		
DISTRIBUTION/AVAILABILITY CODES		
		SPECIAL
A		

S N 0102- LF- 014- 6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

1. Introduction

Given bivariate data $(\,(X_i, Y_i)\; i = 1,\ldots,n\,)$ one would like a technique for estimating the regression function $r(x) = E(Y|X)$, which is applicable for a wide range of functions, $r(x)$. Further, one would like a technique with good precision over the full range of values for X.

Techniques of quantile regression based on the approach developed by Parzen (1977), and discussed by Carmichael (1978), will be compared to each other and to Benedetti (1975) in the case where X is restricted to the unit interval. Further, the quantile regression techniques will be compared in the case where X is a random variable.

Variations of each of the quantile regression estimators will also be considered in order to determine the most functionally appealing form of the estimators.

2. The Estimators

2.1. $\hat{RQ}1$

Given observations $\{(X_i, Y_i),\ i = 1,\ldots,n\}$ on random variables (X, Y) according to the model

$$Y_i = r(X_i) + \varepsilon_i$$

where r is an unknown function and the errors are i.i.d. random variables with zero mean and finite variance, the following versions of the estimator $\hat{RQ1}(u)$ are considered

(a) $$\frac{1}{nh(n)} \sum_{j=1}^{n} Y_{[j:n]} K\left(\frac{\frac{j-1}{n} - u}{h(n)}\right)$$

(b) $$\frac{1}{nh(n)} \left(\frac{Y_{[1:n]}}{2} K\left(\frac{\frac{0}{n} - u}{h(n)}\right) + \frac{Y_{[n:n]}}{2} K\left(\frac{\frac{n-1}{n} - u}{h(n)}\right) + \sum_{j=2}^{n-1} Y_{[j:n]} K\left(\frac{\frac{j-1}{n} - u}{h(n)}\right)\right)$$

(c) $$\frac{\left(\frac{Y_{[1:n]}}{2} K\left(\frac{0-u}{h(n)}\right) + \frac{Y_{[n:n]}}{2} K\left(\frac{\frac{n-1}{n} - u}{h(n)}\right) + \sum_{j=2}^{n-1} Y_{[j:n]} K\left(\frac{\frac{j-1}{n} - u}{h(n)}\right)\right)}{\left(\frac{K\left(\frac{0-u}{h(n)}\right)}{2} + \frac{K\left(\frac{\frac{n-1}{n} - u}{h(n)}\right)}{2} + \sum_{j=2}^{n-1} K\left(\frac{\frac{j-1}{n} - u}{h(n)}\right)\right)}$$

(d) $$\frac{\sum_{j=1}^{n} Y_{[j:n]} K\left(\frac{\frac{j-1}{n} - u}{h(n)}\right)}{\sum_{j=1}^{n} K\left(\frac{\frac{j-1}{n} - u}{h(n)}\right)}$$

where $Y_{[j:n]}$ denotes the observation associated with $X_{(j)}$ where $X_{(1)} < X_{(2)} < \ldots < X_{(n)}$ and $K(Z)$ is defined as follows:

$$K(Z) = \begin{cases} \frac{15}{16} (1 - Z^2)^2 & |Z| \le 1 \\ 0 & |Z| > 1 \end{cases}$$

2.2. $\hat{RQ2}$

If one interpolates linearly between the successive points $\{(j/n, Y_{[j:n]})\}$ one obtains a new estimator $\hat{RQ2}$. The versions of this estimator are as follows:

(a) $$\frac{1}{nh(n)} \sum_{j=1}^{n} Y_{[j:n]} K\left(\frac{\frac{j-1}{n} - u}{h(n)}\right) + \frac{1}{2nh(n)} \sum_{j=2}^{n} \left(Y_{[j-1:n]} - Y_{[j:n]}\right) K\left(\frac{\frac{j-1}{n} - u}{h(n)}\right)$$

(b) $$\frac{\sum_{j=1}^{n} Y_{[j:n]} K\left(\frac{\frac{j-1}{n} - u}{h(n)}\right)}{\sum_{j=1}^{n} K\left(\frac{\frac{j-1}{n} - u}{h(n)}\right)} + \frac{1}{2nh(n)} \sum_{j=2}^{N} \left(Y_{[j-1:n]} - Y_{[j:n]}\right) K\left(\frac{\frac{j-1}{n} - u}{h(n)}\right)$$

2.3. $\hat{RQ3}$

If instead of interpolating, one integrates a smoothed version of the first differences one obtains $\hat{RQ3}$. The versions of this estimator looked at are:

$$\int_{1/2}^{u} \hat{RQ}_1'(s)\, ds + \hat{RQ}_1(1/2)$$

(a) where $\hat{RQ}_1'(s) = \sum_{j=1}^{n-1} \left(Y_{[j+1:n]} - Y_{[j:n]}\right) \frac{1}{h(n)} K\left(\frac{\frac{j-1}{n} - s}{h(n)}\right)$ and $\hat{RQ}_1(1/2)$ is as defined in 2.1(a)

(b) where $\hat{RQ}_1'(s) = \dfrac{n \sum_{j=1}^{n-1} \left(Y_{[j+1:n]} - Y_{[j:n]}\right) K\left(\frac{\frac{j-1}{n} - s}{h(n)}\right)}{\sum_{j=1}^{n-1} K\left(\frac{\frac{j-1}{n} - s}{h(n)}\right)}$

and $\hat{RQ}_1(1/2)$ is as defined in 2.1 (d) .

3. Comparison of the Estimators

One needs a way of comparing the estimators as to their efficiency over a range of functions. One can compare them for each function based on the criterion of minimizing mean integrated squared error (MISE). For fixed $h(n)$, n, and $r(x)$, the MISE is given by

3.1. $$MISE = E\int_0^1 (\hat{R}(x) - r(x))^2\,dx = \int_0^1 VAR(\hat{R}(x))\,dx + \int_0^1 (Bias)^2\,dx \,.$$

One will find in Appendix A the formulas used in the calculation of MISE for $\hat{RQ1}$ and $\hat{RQ3}$, considered both in the case where X is fixed and where X is a random variable.

The MISE is also calculated over the interval (.25, .75) *since* the estimators perform most poorly near the endpoints.

X Fixed

It was found that for a sample of size n , taking points of the form $\frac{j-1}{n-1}$ $j = 1,\ldots,n$, rather than points of the form $\frac{j}{n}$ $j = 1,\ldots,n$, reduced the MISE.

To accomodate this change in design the estimators must be slightly modified. Therefore the estimators $\hat{RQ1}(a)$ and $\hat{RQ3}(a)$ become

3.2. $$\hat{RQ1} = \frac{1}{nh(n)} \sum_{j=1}^{n} Y_{[j:n]} \, K\left(\frac{\frac{j-1}{n-1} - u}{h(n)}\right) \quad \text{and}$$

3.3. $$\hat{RQ3} = \int_{1/2}^{u} \hat{RQ}_1'(s)\,ds + \hat{RQ}_1(1/2) \quad \text{where}$$

$$\hat{RQ}_1'(s) = \sum_{j=1}^{n-1} \left(Y_{[j+1:n]} - Y_{[j:n]}\right) \frac{1}{h(n)} K\left(\frac{\frac{j-1}{n-1} - s}{h(n)}\right)$$

and $\hat{RQ}_1(1/2)$ is as in 3.2.

Similar changes were also made in the other versions of each of the estimators.

Initially one would like to determine which of the four versions of $\hat{RQ1}$ is functionally the best. Tables 1 through 19 are the values calculated for the MISE for the different estimators, for the different models, around the optimum choice for the parameter $M = 1/h(n)$.

A comparison of tables 1, 11, 12 and 13 indicates that the normalized versions of $\hat{RQ1}$ perform better than the un-normalized versions of $\hat{RQ1}$. Therefore, one needs only look at the normalized versions, $\hat{RQ1}(c)$ and $\hat{RQ1}(d)$.

Looking at Table 13 one finds that $\hat{RQ1}(d)$ was better than $\hat{RQ1}(c)$, although little difference is found, and thus $\hat{RQ1}(d)$ was chosen as the version to be looked at in more detail.

In preliminary simulations $\hat{RQ2}$ did not perform substantially different from $\hat{RQ1}$. $\hat{RQ2}(a)$ did not substantially change $\hat{RQ1}(a)$ and therefore can be discarded just on the basis of the poor performance of $\hat{RQ1}(a)$. In these simulations $\hat{RQ2}(b)$ looked very much like $\hat{RQ1}(d)$ and in fact, based on a comparison of tables 1, 2, and 20 had higher integrated squared BIAS. From looking at the formula for $\hat{RQ2}(b)$ there is good reason to suspect an increase in integrated variance. Therefore, $\hat{RQ2}(b)$ will not be considered further.

It is also necessary to compare the two versions of $\hat{RQ3}$ to determine which of these to study further.

One finds, by comparing Tables 1-3 with Tables 14-19, that normalizing $\hat{RQ3}$ leads to a reduction in integrated squared bias but an increase in integrated variance. In fact, in the case where $r(x)$ is linear, the bias of $\hat{RQ3}(d)$ is exactly 0. This is a very desirable result. Still a choice is necessary to determine which merits further study.

One finds that the integrated squared bias and integrated variance are, for most cases, smaller for $\hat{RQ1}(d)$ than $\hat{RQ3}(a)$. Therefore,

there is nothing to be gained in looking further at $\hat{RQ3}(a)$. Therefore, one needs only compare the two estimators $\hat{RQ3}(b)$ and $\hat{RQ1}(d)$ with each other and with the Benedetti estimator.

Tables 1 through 5 are a comparison of the estimators $\hat{RQ3}(b)$ and $\hat{RQ1}(d)$ for the five models on which they were tested. Benedetti only gives the MISE and does not break it down into squared bias and variance. Therefore, one can only compare the MISE of her estimator with the ones discussed here.

Tables 1a, 2a, and 4a indicate that in all of the models tested $\hat{RQ1}(d)$ performed at least as well as the Benedetti estimator and in fact, performed substantially better for the model $Y = SIN(2\pi x)$.

Now one needs to compare $\hat{RQ1}(d)$ and $\hat{RQ3}(b)$. Tables 1 through 5 provide this comparison. One notes that $\hat{RQ1}(d)$ has lower MISE in all the models. This is due to the increase in integrated variance in $\hat{RQ3}(b)$. In fact, the integrated squared bias of $\hat{RQ3}(b)$ was better in the area of the optimum choice of M except for the model $Y = SIN(2\pi x)$ for a sample size of 50. Clearly in all cases for a given value of M, the bias of $\hat{RQ3}(b)$ is smaller than that of $\hat{RQ1}(d)$.

It should be noted that there was anticipation of an improvement by smoothing first differences. It is clear that in the normalized versions of the estimators the bias is reduced by smoothing first differences, but since the variance increases it is unclear whether $\hat{RQ3}$ is an improvement.

If one compares Tables 11 and 17, one finds that when the estimators are not normalized, smoothing first differences represents a substantial improvement. In fact, by also looking at Table 14, one finds that $\widehat{RQ3}(a)$ has properties for a sample size of 20 equivalent to those of $\widehat{RQ1}(a)$ for a sample size of 50. These indeed represent the improvement that was anticipated.

X: Random Variable

For the case where X is a random variable only the estimators $\widehat{RQ1}(d)$ and $\widehat{RQ3}(b)$ are discussed since they were found to be the most appealing estimators for the case when X is fixed.

It is the belief here that a slight modification of the estimators will improve them. Let $U_{(j)} = F(X_{(j)})$, $j = 1,\ldots,n$. These are the order statistics for a uniform distribution. One can calculate their expectations and, in fact,

$$E[U_{(j)}] = \frac{j}{n+1} .$$

Therefore, it is felt that evaluating the kernel at points of the form $\frac{j}{n+1}$ $j = 1,\ldots,n$ will improve the estimators. Initial observations showed that this might in fact be true. Therefore, the estimators are now written as:

$$\widehat{RQ1}(d) = \frac{\sum_{j=1}^{n} Y_{[j:n]} K\left(\frac{\frac{j}{n+1} - u}{h(n)}\right)}{\sum_{j=1}^{n} K\left(\frac{\frac{j}{n+1} - u}{h(n)}\right)}$$

and

$$\widehat{RQ3}(b) = \widehat{RQ1}(1/2) + (n+1)\int_{1/2}^{u} \frac{\sum_{j=1}^{n-1}\left(Y_{[j+1:n]} - Y_{[j:n]}\right) K\left(\frac{\frac{j}{n+1} - s}{h(n)}\right)}{\sum_{j=1}^{n-1} K\left(\frac{\frac{j}{n+1} - s}{h(n)}\right)}\, ds$$

These are the forms of the estimators which are discussed further.

Three models were looked at for two different sample sizes and a few different error structures.

From Tables 21 and 22 for the linear model $Y = 1 - X$, one finds as before that the bias of $\widehat{RQ3}(b)$ is substantially smaller than than of $\widehat{RQ1}(d)$. Also as before, the variance for $\widehat{RQ1}(d)$ is smaller. However, one finds that provided the variance of the error is small enough the MISE is smaller for $\widehat{RQ3}(d)$ and even when the variance of the error increases the MISE of the two estimators are not very different.

Looking at Tables 23 and 24 one compares the performance of the estimators for the model $Y = \text{SIN}(2\pi X)$. One finds similar results. $\widehat{RQ3}(b)$ has substantially smaller bias and larger variance and again the MISE for the two estimators are roughly of the same order.

In Tables 25 and 26 the two estimators are compared for the model $Y = \sqrt{\frac{\sqrt{3}}{16 - 7\sqrt{3}}}\,(X + 4e^{-X^2/2})$. For this model the results are slightly different. Again $\hat{RQ3}(b)$ shows a very large reduction in bias and an increase in variance but for both sample sizes and both error structures the value of the MISE, at the optimum choice of M , for $\hat{RQ3}(b)$ is substantially smaller than that of $\hat{RQ1}(d)$.

The results here clearly show $\hat{RQ3}(b)$ has better bias properties than $\hat{RQ1}(d)$ and also may indicate that the increase in variance may not be very significant. One might conclude that in fact $\hat{RQ3}(b)$ represents an improvement of $\hat{RQ1}(d)$.

One might like to see this improvement so next these two estimators will be compared in simulations.

Simulations

Simulations were performed to see the differences in the estimators and to determine which performed better. Both the case of X fixed and X a random variable were studied.

The first model examined is $Y = 3 + 2X$ when X is fixed. Figures 1, 1a, and 1b indicate the performance of the three estimators. It is clear that $\hat{RQ2}(b)$ does not perform substantially different from $\hat{RQ1}(d)$ and as indicated before it is in fact worse. This lends credence to the claim that $\hat{RQ2}(b)$ is not an improvement and need not be considered further.

Looking only at the other two estimators, one sees that both perform well between .25 and .75 but at the endpoints $\hat{RQ1}(d)$ veers away from the line while $\hat{RQ3}(b)$ is barely discernable from the true line throughout.

When a random error is introduced in this model, one must re-evaluate the estimators. Here $\hat{RQ1}(d)$ may be said to better estimate the true line but $\hat{RQ3}(b)$ seems to follow the data more closely. This can be seen in the area near one. Here the data seems to swing upward and $\hat{RQ3}(b)$ follows this pattern while $\hat{RQ1}(d)$ stays below. (See figures 2 and 2a.)

Next the trigonometric function $Y = SIN(2\pi X)$ with a random error introduced is looked at. The two estimators are not very different except near the endpoints. When X is near zero, it is unclear which performs better. When X is near one, however, the data follows the true regression more closely and $\hat{RQ3}(b)$ better fits the data. (See figures 3 and 3a.)

The same function is examined in the case where X is distributed as a $U(0, 1)$. Again the *differences in the estimators occur* at the endpoints. When X is near zero one sees that the data is close to the true regression and $\hat{RQ3}(b)$ estimates the regression better in this area. When X is near one, there is a clear upswing in the data which also is followed more closely by $\hat{RQ3}(b)$. (See figures 4 and 4a.)

Finally, the function $Y = 1 - X$ with a random error is discussed when X is distributed $N(0, 1)$. Again, the estimators function alike except when X is near one. Here the data has an obvious downward trend which is clearly followed more closely by $\hat{R}Q3(b)$ than $\hat{R}Q1(d)$. *(See figures 5 and 5a.)*

4. Conclusions

It was anticipated that the estimator $\hat{R}Q3(b)$ would perform better especially near the endpoints. This would seem to be the case. Clearly, the estimators are similar in the middle but $\hat{R}Q3(b)$, while having more variability, substantially improves *the estimation at the* endpoints.

Also it was found that normalizing the estimators improved them. This was, however, more important for $\hat{R}Q1$ than $\hat{R}Q3$ but both were improved.

TABLE 1

$Y = 3 + 2X$

$\sigma^2 = .3$

$N = 20$

$M = \frac{1}{h(n)}$

	1.5	2	2.5	3	3.5
Bias Squared (0, 1)					
$\hat{R}Q1(d)$	.0379	.0149	.0071	.0038	.0023
RQ3(b)	0	0	0	0	0
Variance* (0, 1)					
$\hat{R}Q1(d)$	.0225	.0278	.0332	.0386	.0439
$\hat{R}Q3(b)$	.0453	.0441	.0464	.0504	.0552
MISE* (0, 1)					
$\hat{R}Q1(d)$	.0603	.0428	.0403	.0424	.0462
$\hat{R}Q3(b)$	.0453	.0441	.0464	.0504	.0552
Bias Squared (.25, .75)					
$\hat{R}Q1(d)$	.0027	.0002	.0000	.0000	.0000
RQ3(b)	0	0	0	0	0
Variance* (.25, .75)					
$\hat{R}Q1(d)$	.0094	.0116	.0141	.0168	.0196
RQ3(b)	.0103	.0123	.0156	.0193	.0232
MISE* (.25, .75)					
$\hat{R}Q1(d)$	.0121	.0118	.0141	.0168	.0196
RQ3(b)	.0103	.0123	.0156	.0193	.0232

*These values are for $\sigma^2 = .3$. The integrated variance increases proportionally with σ^2 and for a different value of σ^2 must be adjusted. The MISE must also be adjusted accordingly.

$$\text{MISE} = \int_0^1 (\text{BIAS})^2 + \frac{\sigma^2}{.3} \int_0^1 (\text{VAR}) .$$

TABLE 1a

$Y = 3 + 2X$

$\sigma^2 = .3$

$N = 50$

	$M = \frac{1}{h(n)}$ 2	2.5	3	3.5	4
Bias Squared (0,1)					
$\hat{R}Q1(d)$	.0176	.0088	.0050	.0030	.0020
$\hat{R}Q3(b)$	0	0	0	0	0
Variance* (0,1)					
$\hat{R}Q1(d)$	.0111	.0132	.0153	.0175	.0196
$\hat{R}Q3(b)$	.0303	.0280	.0274	.0276	.0283
MISE* (0,1)					
$\hat{R}Q1(d)$	.0287	.0220	.0203	.0205	.0216
$\hat{R}Q3(b)$	.0303	.0280	.0274	.0276	.0283
Bias Squared (.25,.75)					
$\hat{R}Q1(d)$	.0003	.0000	.0000	.0000	.0000
$\hat{R}Q3(b)$	0	0	0	0	0
Variance* (.25,.75)					
$\hat{R}Q1(d)$	.0045	.0055	.0065	.0076	.0087
$\hat{R}Q3(b)$	.0047	.0057	.0069	.0081	.0094
MISE* (.25,.75)					
$\hat{R}Q1(d)$	.0049	.0055	.0065	.0076	.0087
$\hat{R}Q3(b)$	.0047	.0057	.0069	.0081	.0094

Benedetti's Best

MISE*

(0,1)	.022
(.25, .75)	.006

TABLE 2

$Y = \mathrm{SIN}(2\pi X)$

$\sigma^2 = .3$

$N = 20$

	$M = \frac{1}{h(n)}$ 2.5	3	3.5	4	5
Bias Squared (0,1)					
$\hat{R}Q1(d)$		.0447	.0292	.0196	.0095
$\hat{R}Q3(b)$	.0448	.0243	.0171	.0155	
Variance* (0,1)					
$\hat{R}Q1(d)$		.0386	.0439	.0493	.0601
$\hat{R}Q3(b)$	.0464	.0504	.0552	.0608	
MISE* (0,1)					
$\hat{R}Q1(d)$		.0833	.0732	.0690	.0697
$\hat{R}Q3(b)$	.0912	.0747	.0723	.0763	
Bias Squared (.25,.75)					
$\hat{R}Q1(d)$		.0191	.0110	.0067	.0029
$\hat{R}Q3(b)$	.0263	.0117	.0052	.0024	
Variance* (.25,.75)					
$\hat{R}Q1(d)$		.0168	.0196	.0224	.0280
$\hat{R}Q3(b)$	.0156	.0193	.0232	.0229	
MISE* (.25,.75)					
$\hat{R}Q1(d)$		.0359	.0306	.0291	.0309
$\hat{R}Q3(b)$	.0419	.0310	.0284	.0253	

TABLE 2a

$Y = SIN(2\pi X)$

$\sigma^2 = .3$

$N = 50$

$M = \frac{1}{h(n)}$

	3	3.5	4	5	6	8
Bias Squared (0,1)						
$\hat{R}Q1(d)$			.0224	.0117	.0067	.0026
$\hat{R}Q3(b)$	.0258	.0137	.0078	.0040		
Variance* (0,1)						
$\hat{R}Q1(d)$			.0196	.0239	.0282	.0367
$\hat{R}Q3(b)$	.0274	.0276	.0283	.0386		
MISE* (0,1)						
$\hat{R}Q1(d)$			.0420	.0356	.0348	.0393
$\hat{R}Q3(b)$	.0532	.0413	.0361	.0426		
Bias Squared (.25,.75)						
$\hat{R}Q1(d)$			.0067	.0029	.0014	.0004
$\hat{R}Q3(b)$	.0155	.0081	.0044	.0014		
Variance* (.25,.75)						
$\hat{R}Q1(d)$			.0087	.0109	.0130	.0174
$\hat{R}Q3(b)$	.0069	.0081	.0094	.0119		
MISE* (.25,.75)						
$\hat{R}Q1(d)$			.0154	.0138	.0145	.0178
$\hat{R}Q3(b)$	.0224	.0162	.0138	.0133		

Benedetti's Best

MISE*

(0, 1)	.073
(.25, .75)	.029

TABLE 3

$Y = -X|2$

$\sigma^2 = .3$

$N = 20$

$M = 1/h(n)$

	1	1.5	2	2.5	3
Bias Squared (0,1)					
$\hat{R}Q1(d)$	.0076	.0024	.0009	.0004	.0002
$\hat{R}Q3(b)$	0	0	0	0	0
Variance* (0,1)					
$\hat{R}Q1(d)$	.0174	.0225	.0278	.0332	.0386
$\hat{R}Q3(b)$	.0492	.0453	.0441	.0464	.0504
MISE* (0,1)					
$\hat{R}Q1(d)$	.0250	.0249	.0288	.0336	.0388
$\hat{R}Q3(b)$	.0492	.0453	.0441	.0464	.0504
Bias Squared (.25,.75)					
$\hat{R}Q1(d)$	.0010	.0002	.0000	.0000	.0000
$\hat{R}Q3(b)$	0	0	0	0	0
Variance* (.25,.75)					
$\hat{R}Q1(d)$	.0079	.0094	.0116	.0141	.0168
$\hat{R}Q3(b)$	.0111	.0103	.0123	.0156	.0193
MISE* (.25,.75)					
$\hat{R}Q1(d)$	.0089	.0096	.0116	.0141	.0168
$\hat{R}Q3(b)$	.0111	.0103	.0123	.0156	.0193

TABLE 3a

$Y = -X|2$
$\sigma^2 = .3$
$N = 50$

$M = \frac{1}{h(n)}$

	1	1.5	2	2.5	3	3.5	4
Bias Squared (0, 1)							
$\hat{R}Q1(d)$	.0081	.0027	.0011	.0005	.0003		
$\hat{R}Q3(b)$			0	0	0	0	0
Variance* (0, 1)							
$\hat{R}Q1(d)$	.0069	.0089	.0111	.0132	.0153		
$\hat{R}Q3(b)$			.0303	.0280	.0274	.0276	.0283
MISE* (0, 1)							
$\hat{R}Q1(d)$	.0150	.0116	.0122	.0137	.0156		
$\hat{R}Q3(b)$			.0303	.0280	.0274	.0276	.0283
Bias Squared (.25, .75)							
$\hat{R}Q1(d)$	.0011	.0002	.0000	.0000	.0000		
$\hat{R}Q3(b)$			0	0	0	0	0
Variance* (.25, .75)							
$\hat{R}Q1(d)$	.0031	.0037	.0045	.0055	.0065		
$\hat{R}Q3(b)$			.0047	.0057	.0069	.0081	.0094
MISE* (.25, .75)							
$\hat{R}Q1(d)$	.0042	.0039	.0046	.0055	.0065		
$\hat{R}Q3(b)$			.0047	.0057	.0069	.0081	.0094

TABLE 4

$Y = 1 - X$
$\sigma^2 = .3$
$N = 20$

$M = 1/h(n)$

	1.5	2	2.5	3
Bias Squared (0, 1)				
$\hat{R}Q1(d)$	.0095	.0037	.0018	.0010
$\hat{R}Q3(b)$	0	0	0	0
Variance* (0, 1)				
$\hat{R}Q1(d)$	.0225	.0278	.0332	.0386
$\hat{R}Q3(b)$	.0453	.0441	.0464	.0504
MISE* (0, 1)				
$\hat{R}Q1(d)$	.0320	.0316	.0350	.0395
$\hat{R}Q3(b)$	.0453	.0441	.0464	.0504
Bias Squared (.25, .75)				
$\hat{R}Q1(d)$	.0007	.0001	.0000	.0000
$\hat{R}Q3(b)$	0	0	0	0
Variance* (.25, .75)				
$\hat{R}Q1(d)$	.0094	.0116	.0141	.0168
$\hat{R}Q3(b)$	.0103	.0123	.0156	.0193
MISE* (.25, .75)				
$\hat{R}Q1(d)$	.0101	.0116	.0141	.0168
$\hat{R}Q3(b)$	.0103	.0123	.0156	.0193

TABLE 4a

$Y = 1 - X$

$\sigma^2 = .3$

$N = 50$

$M = 1/h(n)$

	1.5	2	2.5	3	3.5	4
Bias Squared (0, 1)						
$\hat{R}Q1(d)$	.0107	.0044	.0022	.0012	.0008	
$\hat{R}Q3(b)$		0	0	0	0	0
Variance* (0, 1)						
$\hat{R}Q1(d)$	.0089	.0111	.0132	.0153	.0175	
$\hat{R}Q3(b)$		.0303	.0280	.0274	.0276	.0283
MISE* (0, 1)						
$\hat{R}Q1(d)$	.0196	.0155	.0154	.0166	.0182	
$\hat{R}Q3(b)$		.0303	.0280	.0274	.0276	.0183
Bias Squared (.25, .75)						
$\hat{R}Q1(d)$	.0008	.0001	.0000	.0000	.0000	
$\hat{R}Q3(b)$		0	0	0	0	0
Variance* (.25, .75)						
$\hat{R}Q1(d)$	.0037	.0045	.0055	.0065	.0076	
$\hat{R}Q3(b)$		.0047	.0057	.0069	.0081	.0094
MISE* (.25, .75)						
$\hat{R}Q1(d)$	.0045	.0046	.0055	.0065	.0076	
$\hat{R}Q3(b)$		.0047	.0057	.0069	.0081	.0094

Benedetti's Best

MISE*

(0, 1)	.016
(.25, .75)	.005

TABLE 5

$Y = \sqrt{\frac{3}{7}}\,(X + 4e^{-X^2/2})$

$\sigma^2 = .3$

$N = 20$

$M = \frac{1}{h(n)}$

	1	1.5	2	2.5	3
Bias Squared (0, 1)					
$\hat{R}Q1(d)$	.0092	.0039	.0018	.0010	
$\hat{R}Q3(b)$		.0011	.0004	.0002	.0001
Variance* (0, 1)					
$\hat{R}Q1(d)$	.0174	.0225	.0278	.0332	
$\hat{R}Q3(b)$		.0453	.0441	.0464	.0504
MISE* (0, 1)					
$\hat{R}Q1(d)$	.0266	.0264	.0297	.0342	
$\hat{R}Q3(b)$		.0464	.0445	.0466	.0505
Bias Squared (.25, .75)					
$\hat{R}Q1(d)$	.0024	.0010	.0004	.0002	
$\hat{R}Q3(b)$		.0009	.0003	.0001	.0000
Variance* (.25, .75)					
$\hat{R}Q1(d)$	.0079	.0094	.0116	.01[illegible]1	
$\hat{R}Q3(b)$		.0103	.0123	.0156	.0193
MISE* (.25, .75)					
$\hat{R}Q1(d)$	.0103	.0104	.0120	.0143	
$\hat{R}Q3(b)$		.0112	.0126	.0157	.0193

TABLE 5a

$Y = \sqrt{\frac{3}{7}}(X + 4e^{-x^2/2})$

$\sigma^2 = .3$

$N = 50$

	M = 1/h(n)				
	1.5	2	2.5	3	3.5
Bias Squared (0, 1)					
$\hat{R}Q1(d)$	.0043	.0021	.0011	.0007	
$\hat{R}Q3(b)$		.0005	.0002	.0001	.0000
Variance* (0, 1)					
$\hat{R}Q1(d)$	.0089	.0111	.0132	.0153	
$\hat{R}Q3(b)$		.0303	.0280	.0274	.0276
MISE* (0, 1)					
$\hat{R}Q1(d)$	.0132	.0131	.0143	.0160	
$\hat{R}Q3(b)$		.0308	.0282	.0275	.0276
Bias Squared (.25, .75)					
$\hat{R}Q1(d)$	.0010	.0004	.0002	.0001	
$\hat{R}Q3(b)$		.0003	.0002	.0001	.0000
Variance* (.25, .75)					
$\hat{R}Q1(d)$	.0037	.0045	.0055	.0065	
$\hat{R}Q3(b)$		.0047	.0057	.0069	.0081
MISE* (.25, .75)					
$\hat{R}Q1(d)$	.0047	.0049	.0057	.0066	
$\hat{R}Q3(b)$		.0050	.0059	.0070	.0081

TABLE 6

$Y = 3 + 2X$

$\sigma^2 = .3$

$N = 100$

	M = 1/h(n)			
$\hat{R}Q1(d)$	3	3.5	4	5
Bias Squared				
(0, 1)	.0054	.0033	.0022	.0011
(.25, .75)	.0000	.0000	.0000	.0000
Variance*				
(0, 1)	.0076	.0087	.0098	.0119
(.25, .75)	.0032	.0038	.0043	.0054
MISE*				
(0, 1)	.0130	.0121	.0120	.0130
(.25, .75)	.0032	.0038	.0043	.0054

TABLE 7

$Y = \sin 2\pi X$

$\sigma^2 = .3$

$N = 100$

	M = 1/h(n)			
$\hat{R}Q1(d)$	5	6	8	10
Bias Squared				
(0, 1)	.0126	.0073	.0030	.0015
(.25, .75)	.0029	.0014	.0004	.0002
Variance*				
(0, 1)	.0119	.0141	.0183	.0226
(.25, .75)	.0054	.0065	.0086	.0108
MISE*				
(0, 1)	.0245	.0214	.0213	.0241
(.25, .75)	.0083	.0079	.0091	.0109

TABLE 8

$Y = 1 - X$
$\sigma^2 = .3$
$N = 100$

$\hat{RQ}1(d)$	M = 1/h(n)			
	2	2.5	3	3.5
Bias Squared				
(0,1)	.0046	.0023	.0013	.0008
(.25,.75)	.0001	.0000	.0000	.0000
Variance*				
(0,1)	.0055	.0066	.0076	.0087
(.25,.75)	.0023	.0027	.0032	.0038
MISE*				
(0,1)	.0102	.0089	.0090	.0096
(.25,.75)	.0024	.0027	.0032	.0038

TABLE 9

$Y = -X|2$
$\sigma^2 = .3$
$N = 100$

$\hat{RQ}1(d)$	M = 1/h(n)			
	1.5	2	2.5	3
Bias Squared				
(0,1)	.0028	.0012	.0006	.0003
(.25,.75)	.0002	.0000	.0000	.0000
Variance*				
(0,1)	.0044	.0055	.0066	.0076
(.25,.75)	.0018	.0023	.0027	.0032
MISE*				
(0,1)	.0072	.0067	.0072	.0080
(.25,.75)	.0021	.0023	.0027	.0032

TABLE 10

$Y = \sqrt{3/7}\,(X + 4e^{-x^2/2})$
$\sigma^2 = .3$
$N = 100$

$\hat{RQ}1(d)$	M = 1/h(n)			
	1.5	2	2.5	3
Bias Squared				
(0,1)	.0044	.0021	.0012	.0007
(.25,.75)	.0010	.0004	.0002	.0001
Variance*				
(0,1)	.0044	.0055	.0066	.0076
(.25,.75)	.0018	.0023	.0027	.0032
MISE*				
(0,1)	.0088	.0077	.0078	.0084
(.25,.75)	.0028	.0026	.0029	.0033

TABLE 11

$Y = 3 + 2X$
$\sigma^2 = .3$
$N = 50$

$\hat{RQ}1(a)$	M = 1/h(n)				
	12	16	20	24	28
Bias Squared					
(0,1)	.0818	.0498	.0319	.0226	.0191
(.25,.75)	.0032	.0032	.0032	.0033	.0051
Variance*					
(0,1)	.0497	.0665	.0834	.1002	.1171
(.25,.75)	.0250	.0334	.0417	.0501	.0584
MISE*					
(0,1)	.1314	.1163	.1153	.1228	.13t2
(.25,.75)	.0282	.0366	.0449	.0534	.0635

TABLE 12

$Y = 3 + 2X$

$\sigma^2 = .3$

$N = 50$

$\hat{R}Q1(b)$	M = 1/h(n) 16	20	24
Bias Squared			
(0, 1)	.1072	.0875	.0768
(.25, .75)	.0000	.0000	.0002
Variance*			
(0, 1)	.0682	.0855	.1028
(.25, .75)	.0348	.0434	.0521
MISE			
(0, 1)	.1754	.1730	.1796
(.25, .75)	.0348	.0434	.0523

TABLE 13

$Y = 3 + 2X$

$\sigma^2 = .3$

$N = 50$

	M = 1/h(n) 16	20	24
Bias Squared (0, 1)			
$\hat{R}Q1(c)$	.0000	.0000	.0000
$\hat{R}Q1(d)$	.0000	.0000	.0000
Variance* (0, 1)			
$\hat{R}Q1(c)$	.0715	.0888	.1061
$\hat{R}Q1(d)$	.0711	.0883	.1055
MISE* (0, 1)			
$\hat{R}Q1(c)$	.0715	.0888	.1061
$\hat{R}Q1(d)$	.0711	.0883	.1055
Bias Squared (.25, .75)			
$\hat{R}Q1(c)$	.0000	.0000	.0000
$\hat{R}Q1(d)$	.0000	.0000	.0000
Variance* (.25, .75)			
$\hat{R}Q1(c)$	.0348	.0434	.0521
$\hat{R}Q1(d)$	.0348	.0434	.0521
MISE* (.25, .75)			
$\hat{R}Q1(c)$	.0348	.0434	.0521
$\hat{R}Q1(d)$	.0348	.0434	.0521

TABLE 14

$Y = 3 + 2X$

$\sigma^2 = .3$

$N = 20$

$\hat{R}Q3(a)$	M = 1/h(n) 1.5	2	2.5	3
Bias Squared				
(0, 1)	.0871	.0465	.0446	.0424
(.25, .75)	.0380	.0202	.0201	.0196
Variance*				
(0, 1)	.0338	.0362	.0384	.0418
(.25, .75)	.0088	.0105	.0128	.0153
MISE*				
(0, 1)	.1210	.0827	.0831	.0843
(.25, .75)	.0468	.0307	.0329	.0348

TABLE 15

$Y = \text{SIN } 2\pi X$

$\sigma^2 = .3$

$N = 20$

$\hat{R}Q3(a)$	M = 1/h(n) 3	3.5	4	5
Bias Squared				
(0, 1)	.0551	.0445	.0393	.0356
(.25, .75)	.0225	.0144	.0101	.0062
Variance*				
(0, 1)	.0418	.0458	.0502	.0599
(.25, .75)	.0153	.0179	.0206	.0261
MISE*				
(0, 1)	.0969	.0904	.0895	.0954
(.25, .75)	.0378	.0324	.0307	.0323

TABLE 16

$Y = -X|2$
$\sigma^2 = .3$
$N = 20$

$\hat{RQ3}(a)$	M = 1/h(n): 1.5	2	2.5	3
Bias Squared				
(0, 1)	.0009	.0004	.0003	.0003
(.25, .75)	.0002	.0001	.0001	.0001
Variance*				
(0, 1)	.0338	.0362	.0384	.0418
(.25, .75)	.0088	.0105	.0128	.0153
MISE*				
(0, 1)	.0347	.0366	.0388	.0421
(.25, .75)	.0090	.0106	.0128	.0154

TABLE 17

$Y = 3 + 2X$
$\sigma^2 = .3$
$N = 50$

$\hat{RQ3}(a)$	M = 1/h(n): 2.5	3	3.5	4
Bias Squared				
(0, 1)	.0085	.0077	.0073	.0070
(.25, .75)	.0032	.0032	.0032	.0032
Variance*				
(0, 1)	.0225	.0227	.0234	.0245
(.25, .75)	.0053	.0063	.0073	.0084
MISE*				
(0, 1)	.0309	.0303	.0307	.0315
(.25, .75)	.0085	.0094	.0105	.0115

TABLE 18

$Y = \sin 2\pi X$
$\sigma^2 = .3$
$N = 50$

$\hat{RQ3}(a)$	M = 1/h(n): 4	5	6	8
Bias Squared				
(0, 1)	.0183	.0112	.0084	.0065
(.25, .75)	.0072	.0034	.0019	.0009
Variance*				
(0, 1)	.0245	.0273	.0306	.0381
(.25, .75)	.0084	.0105	.0126	.0169
MISE*				
(0, 1)	.0427	.0385	.0390	.0446
(.25, .75)	.0156	.0139	.0145	.0179

TABLE 19

$Y = -X|2$
$\sigma^2 = .3$
$N = 50$

$\hat{RQ3}(a)$	M = 1/h(n): 2	2.5	3	3.5
Bias Squared				
(0, 1)	.0002	.0001	.0001	.0001
(.25, .75)	.0000	.0000	.0000	.0000
Variance*				
(0, 1)	.0232	.0225	.0227	.0234
(.25, .75)	.0045	.0053	.0063	.0073
MISE*				
(0, 1)	.0235	.0226	.0227	.0235
(.25, .75)	.0045	.0053	.0063	.0073

TABLE 20

N = 50

$\hat{R}Q2(b)$	M = 1/h(n): 2	3	4	5
Y = 3 + 2X :				
Bias Squared				
(0, 1)	.0180	.0053	.0024	.0014
(.25, .75)	.0006	.0002	.0002	.0002
Y = SIN(2πX)				
Bias Squared				
(0, 1)	.1157	.0490	.0240	.0135
(.25, .75)	.0629	.0199	.0077	.0039

TABLE 21

Y = 1 - X

X ~ N(0, 1)

N = 20

	M = 1/h(n): 1.5	2	2.5	3	3.5
Bias Squared $(\frac{1}{302}, \frac{301}{302})$					
$\hat{R}Q1(d)$			.0813	.0510	.0463
$\hat{R}Q3(b)$	.0379	.0241	.0204	.0183	
$\epsilon \sim N(0, 3)$					
Variance $(\frac{1}{302}, \frac{301}{302})$					
$\hat{R}Q1(d)$			.1072	.1159	.1243
$\hat{R}Q3(b)$	.1396	.1379	.1395	.1433	
MISE $(\frac{1}{302}, \frac{301}{302})$					
$\hat{R}Q1(d)$			.1885	.1769	.1726
$\hat{R}Q3(b)$	.1775	.1620	.1599	.1616	
$\epsilon \sim N(0, 1)$					
Variance $(\frac{1}{302}, \frac{301}{302})$					
$\hat{R}Q1(d)$			.1736	.1936	.2137
$\hat{R}Q3(b)$	.2373	.2338	.2390	.2496	
MISE $(\frac{1}{302}, \frac{301}{302})$					
$\hat{R}Q1(d)$			.2545	.2536	.2620
$\hat{R}Q3(b)$	.2752	.2579	.2679		
$\epsilon \sim$ Double Exp. $\sigma^2 = .3$					
Variance $(\frac{1}{302}, \frac{301}{302})$					
$\hat{R}Q1(d)$			.1167	.1267	.1358
$\hat{R}Q3(b)$	.1416	.1442	.1487	.1550	
MISE $(\frac{1}{302}, \frac{301}{302})$					
$\hat{R}Q1(d)$			.1980	.1877	.1841
$\hat{R}Q3(b)$	.1795	.1683	.1691	.1733	

TABLE 22

Y = 1 - X

X ~ N(0, 1)

N = 50

	M					
	2.5	3	3.5	4	5	6
Bias Squared $(\frac{1}{302}, \frac{301}{302})$						
$\hat{R}Q1(d)$			.0437	.0353	.0283	.0221
$\hat{R}Q3(b)$	.0198	.0144	.0110	.0102		
$\epsilon \sim N(0, .3)$						
Variance $(\frac{1}{302}, \frac{301}{302})$						
$\hat{R}Q1(d)$			.0604	.0638	.0702	.0763
$\hat{R}Q3(b)$	.0754	.0755	.0766	.0782		
MISE $(\frac{1}{302}, \frac{301}{302})$						
$\hat{R}Q1(d)$			.1041	.0991	.0985	.0984
$\hat{R}Q3(b)$	.0952	.0899	.0876	.0884		
$\epsilon \sim N(0, 1)$						
Variance $(\frac{1}{302}, \frac{301}{302})$						
$\hat{R}Q1(d)$			.1030	.1113	.1277	.1440
$\hat{R}Q3(b)$	.1398	.1391	.1414	.1454		
MISE $(\frac{1}{302}, \frac{301}{302})$						
$\hat{R}Q1(d)$			.1467	.1466	.1560	.1661
$\hat{R}Q3(b)$	.1596	.1535	.1524	.1556		

TABLE 23

Y = SIN 2πX

X ~ U(0, 1)

N = 20

	M = 1/h(n)				
	2.5	3	3.5	4	5
Bias Squared (0, 1)					
$\hat{R}Q1(d)$		.0772	.0619	.0515	.0385
$\hat{R}Q3(b)$	.0805	.0557	.0445	.0373	
$\epsilon \sim N(0, .3)$					
Variance (0, 1)					
$\hat{R}Q1(d)$		.1002	.1140	.1265	.1481
$\hat{R}Q3(b)$	.1066	.1220	.1370	.1513	
MISE (0, 1)					
$\hat{R}Q1(d)$		.1774	.1759	.1780	.1866
$\hat{R}Q3(b)$	.1871	.1777	.1815	.1886	
$\epsilon \sim U(-1, 1)$					
Variance (0, 1)					
$\hat{R}Q1(d)$		.1025	.1157	.1278	.1493
$\hat{R}Q3(b)$	.1244	.1380	.1506	.1623	
MISE (0, 1)					
$\hat{R}Q1(d)$		.1797	.1776	.1793	.1878
$\hat{R}Q3(b)$	.2049	.1937	.1951	.1996	

TABLE 24

$Y = \mathrm{SIN}(2\pi X)$

$X \sim U(0,1)$

$N = 50$

	M = 1/h(n)				
	3.5	4	5	6	8
Bias Squared (0, 1)					
$R\hat{Q}1(d)$		.0354	.0228	.0159	.0094
$R\hat{Q}3(b)$	.0256	.0180	.0127	.0103	
$\epsilon \sim N(0, .3)$					
Variance (0, 1)					
$R\hat{Q}1(d)$		.0497	.0585	.0658	.0782
$R\hat{Q}3(b)$	.0545	.0583	.0621	.0771	
MISE (0, 1)					
$R\hat{Q}1(d)$		.0851	.0813	.0817	.0876
$R\hat{Q}3(b)$	.0801	.0763	.0748	.0884	
$\epsilon \sim U(-1, 1)$					
Variance (0, 1)					
$R\hat{Q}1(d)$		.0452	.0539	.0615	.0747
$R\hat{Q}3(b)$	.0555	.0582	.0612	.0815	
MISE (0, 1)					
$R\hat{Q}1(d)$		.0806	.0767	.0774	.0841
$R\hat{Q}3(b)$	.0811	.0762	.0739	.0915	

TABLE 25

$Y = \sqrt{\frac{\sqrt{3}}{16-7\sqrt{3}}}\ (X + 4e^{-x^2/2})$

$X \sim N(0,1)$

$N = 20$

	M = 1/h(n)					
	2.5	3	3.5	4	5	6
Bias Squared $(\frac{1}{302}, \frac{301}{302})$						
$R\hat{Q}1(d)$		.1766	.1428	.1209	.0916	.0738
$RQ3(b)$	.0941	.0731	.0619	.0551	.0473	
$\epsilon \sim N(0, .3)$						
Variance $(\frac{1}{302}, \frac{301}{302})$						
$R\hat{Q}1(d)$		.1844	.1985	.2105	.2313	.2495
$R\hat{Q}3(b)$	.1951	.2064	.2172	.2280	.2485	
MISE $(\frac{1}{302}, \frac{301}{302})$						
$R\hat{Q}1(d)$		.3610	.3423	.3314	.3229	.3233
$R\hat{Q}3(b)$	.2892	.2795	.2791	.2831	.2958	
$\epsilon \sim N(0, 1)$						
Variance $(\frac{1}{302}, \frac{301}{302})$						
$R\hat{Q}1(d)$		.2730	.2993	.3235	.3693	.4128
$R\hat{Q}3(b)$	.3009	.3198	.3379	.3590	.4037	
MISE $(\frac{1}{302}, \frac{301}{302})$						
$R\hat{Q}1(d)$		.4496	.4431	.4444	.4609	.4866
$R\hat{Q}3(b)$	.3950	.3920	.3998	.4141	.4510	

TABLE 26

$$Y = \sqrt{\frac{\sqrt{3}}{16-7\sqrt{3}}}\,(X + 4e^{-x^2/x})$$

$X \sim N(0,1)$

$N = 50$

$M = 1/h(n)$

Bias Squared $(\frac{1}{302}, \frac{301}{302})$						
$\hat{R}Q1(d)$			.0757	.0478	.0330	.0197
$\hat{R}Q3(b)$	.0297	.0203	.0152	.0075		
$\epsilon \sim N(0,.3)$						
Variance $(\frac{1}{302}, \frac{301}{302})$						
$\hat{R}Q1(d)$			.0804	.0845	.0973	.1108
$\hat{R}Q3(b)$	.0829	.0866	.0903	.1129		
MISE $(\frac{1}{302}, \frac{301}{302})$						
$\hat{R}Q1(d)$			.1561	.1323	.1303	.1305
$\hat{R}Q3(b)$	.1126	.1069	.1055	.1204		
$\epsilon \sim N(0,1)$						
Variance $(\frac{1}{302}, \frac{301}{302})$						
$\hat{R}Q1(d)$			.1195	.1387	.1569	.1910
$\hat{R}Q3(b)$	.1424	.1466	.1521	.1952		
MISE $(\frac{1}{302}, \frac{301}{302})$						
$\hat{R}Q1(d)$			.1952	.1865	.1899	.2107
$\hat{R}Q3(b)$	.1721	.1669	.1673	.2027		

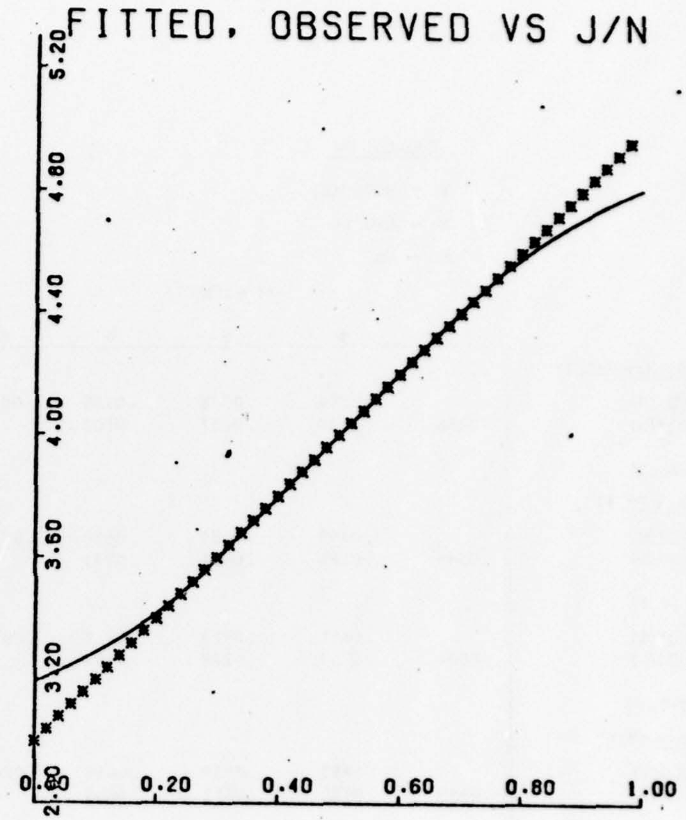


Figure 1.
$\hat{R}Q1(d)$

Y = 3 + 2X
X : fixed
N = 50
M = 3

KEY:

* DATA
Y TRUE REGRESSION
ESTIMATOR

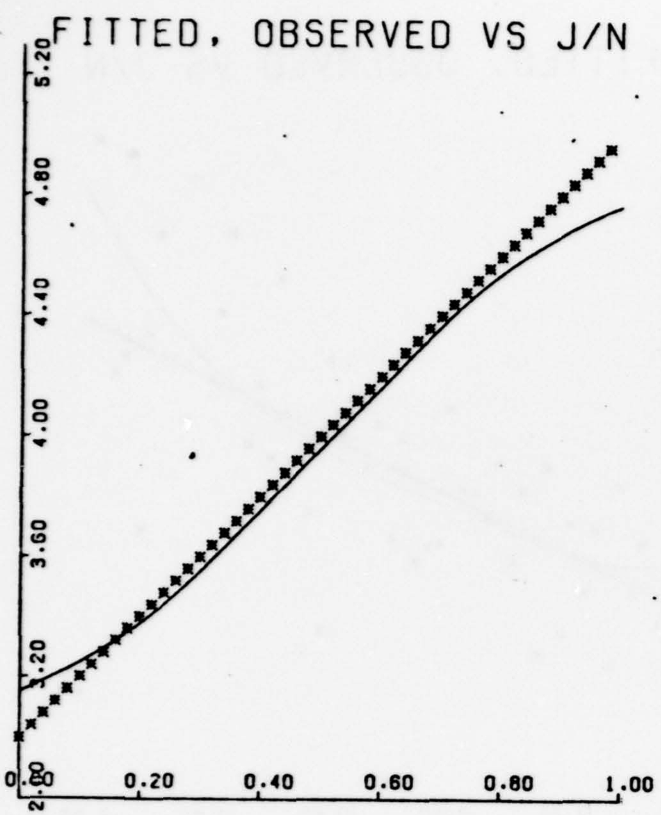


Figure 1a.

$\hat{RQ2}$ (b)

Y = 3 + 2X
X : fixed
N = 50
M = 3

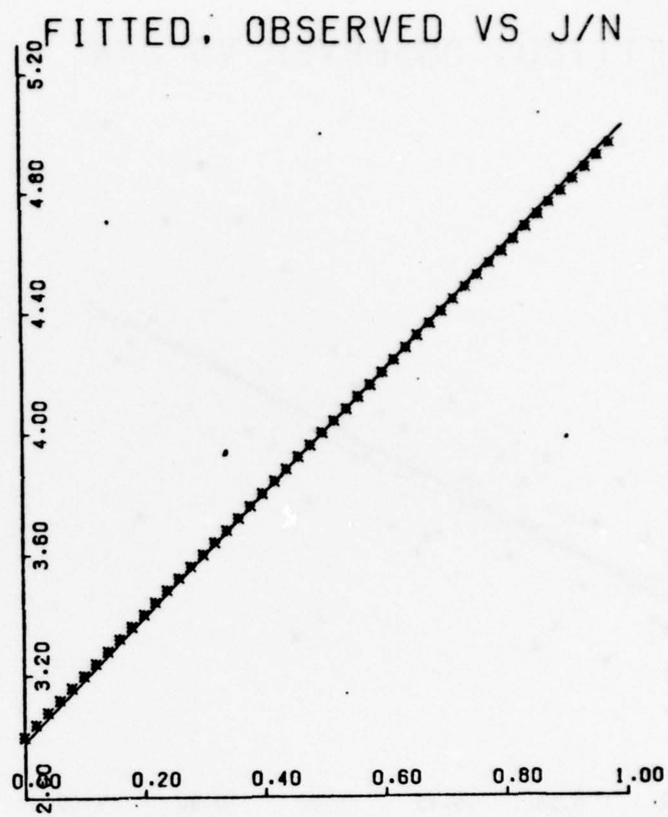


Figure 1b.

$\hat{RQ3}$ (d)

Y = 3 + 2X
X : fixed
N = 50
M = 3

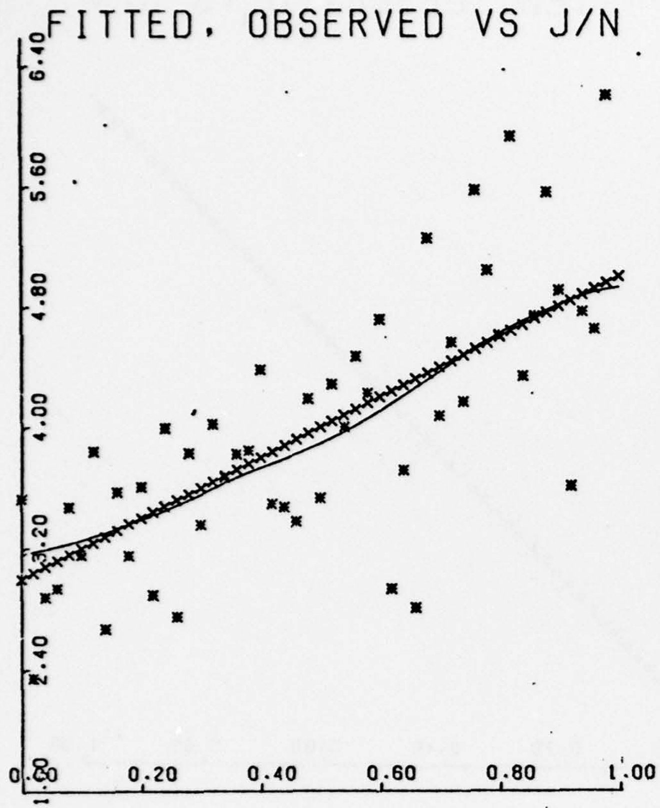


Figure 2.

$\hat{RQ1}$ (d)

$Y = 3 + 2X + \epsilon$
$\epsilon \sim N(0, .3)$
X : fixed
N = 50
M = .3

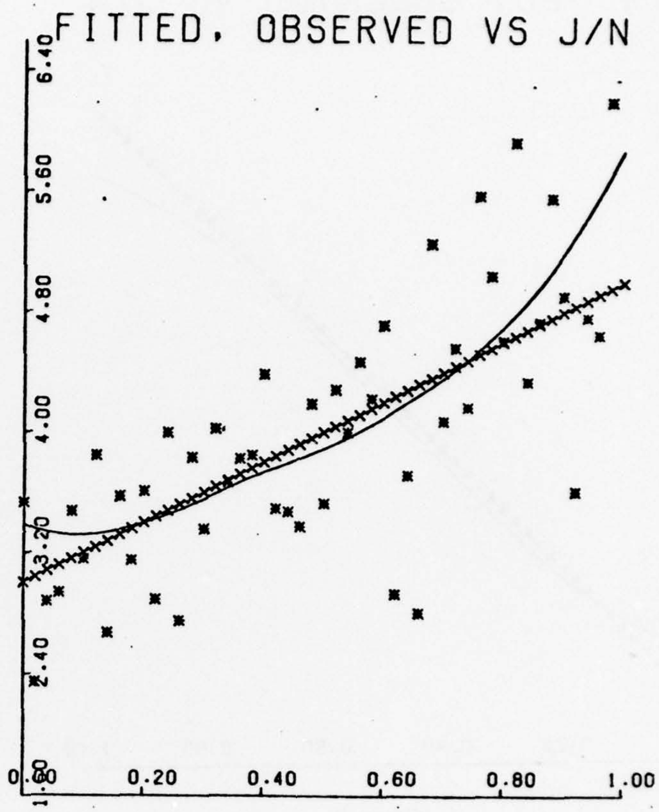


Figure 2a.

$\hat{RQ3}$ (b)

$Y = 3 + 2X + \epsilon$
$\epsilon \sim N(0, .3)$
X : fixed
N = 50
M = 3

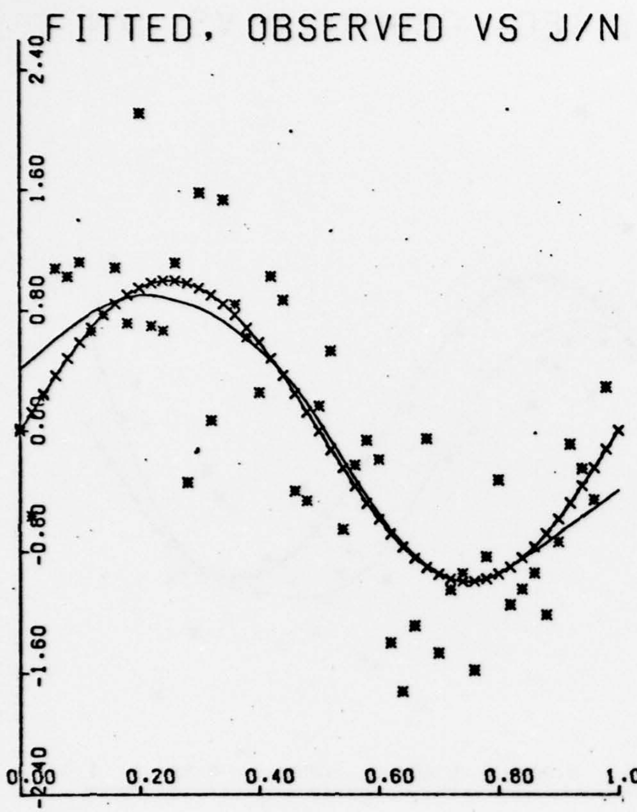


Figure 3.

$\hat{RQ1}$ (d)

$Y = SIN(2\pi X) + \epsilon$
$\epsilon \sim N(0, .3)$
X : fixed
N = 50
M = 5

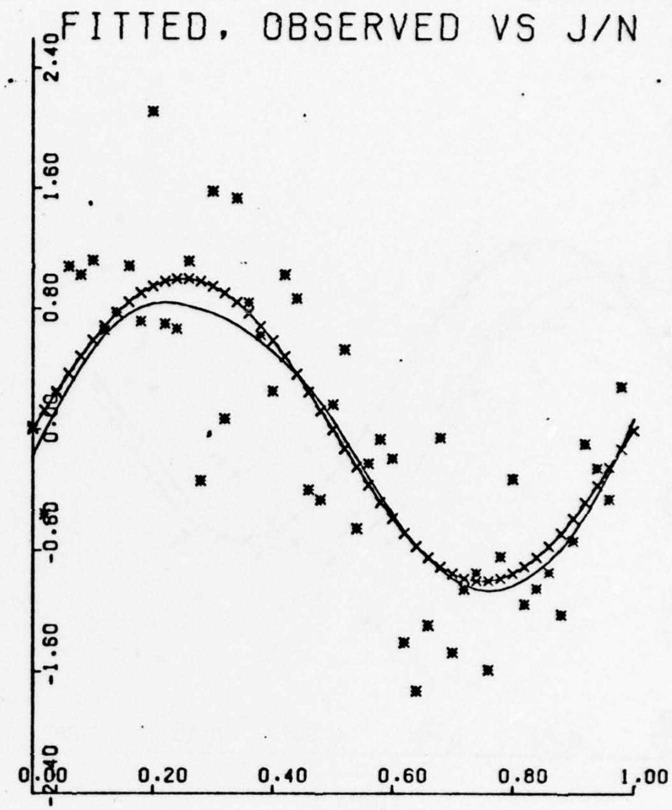


Figure 3a.

$\hat{RQ3}$ (b)

$Y = SIN(2\pi X) + \epsilon$
$\epsilon \sim N(0, .3)$
X : fixed
N = 50
M = 5

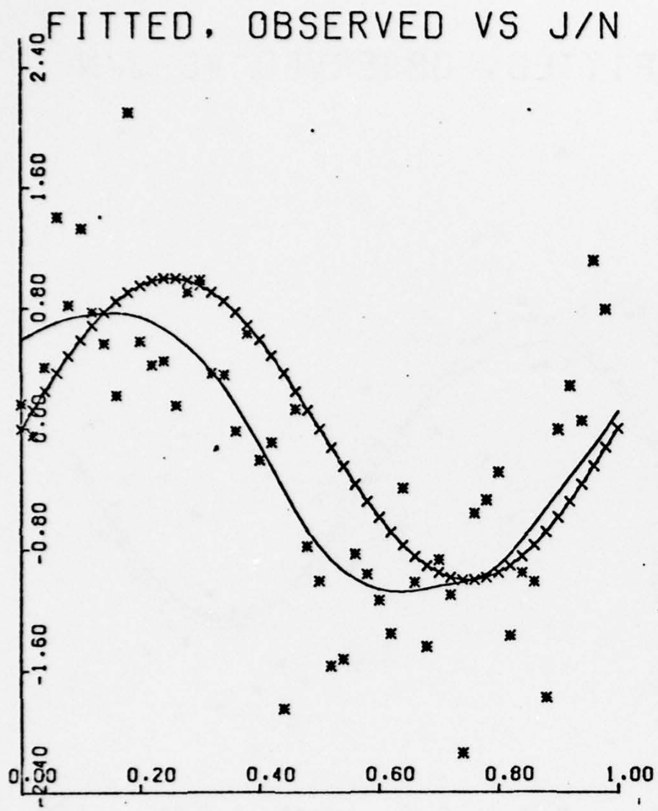


Figure 4.

$\hat{RQ1}$ (d)

$Y = SIN(2\pi X) + \epsilon$
$\epsilon \sim N(0, .3)$
$X \sim U(0, 1)$
$N = 50$
$M = 5$

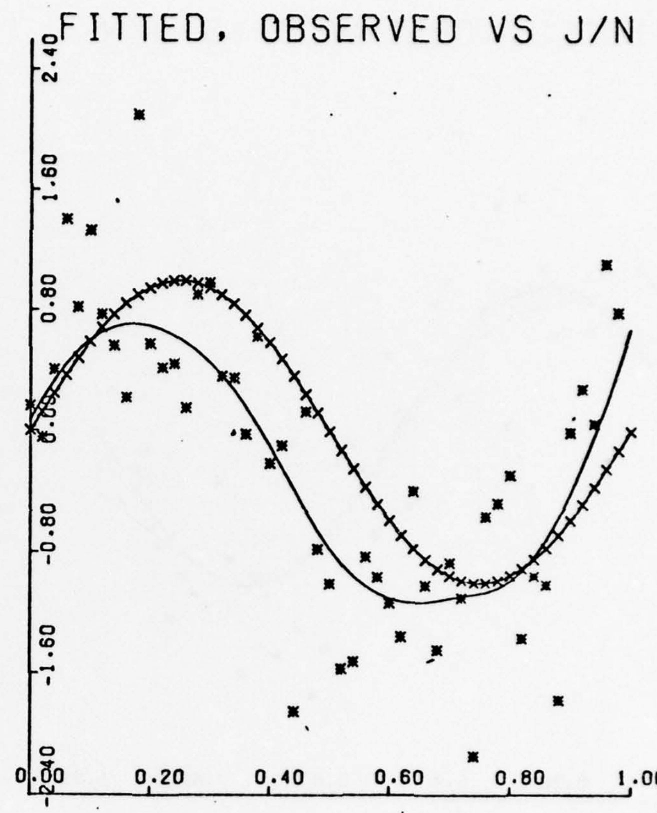


Figure 4a.

$\hat{RQ3}$ (b)

$Y = SIN(2\pi X) + \epsilon$
$\epsilon \sim N(0, .3)$
$X \sim U(0, 1)$
$N = 50$
$M = 5$

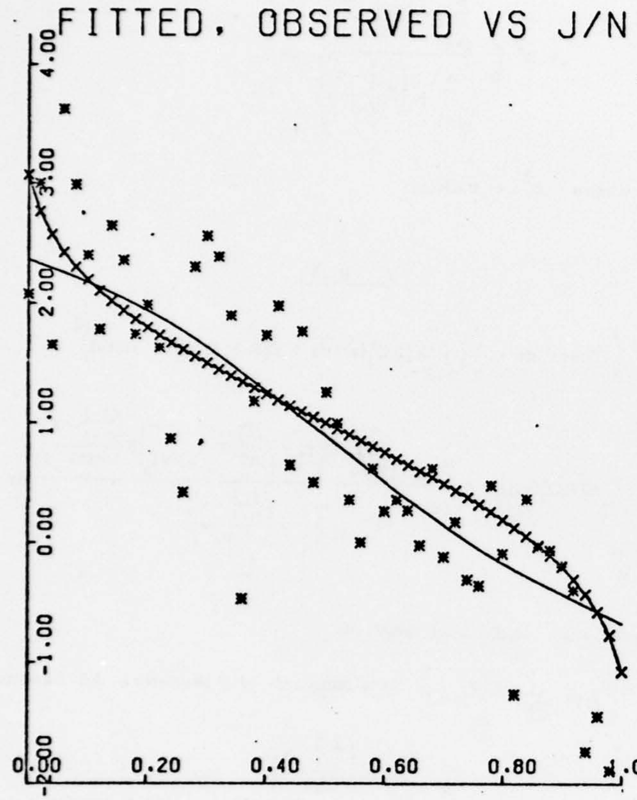


Figure 5.

$\hat{RQ1}$ (d)

$Y = 1 - X + \epsilon$
$X \sim N(0,1)$
$\epsilon \sim N(0,.3)$
$N = 50$
$M = 3$

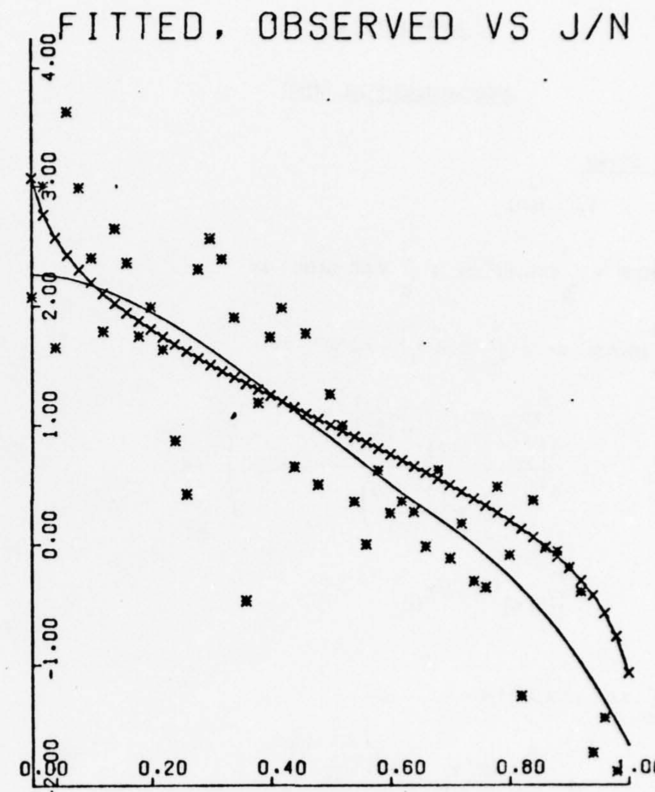


Figure 5a.

$\hat{RQ3}$ (b)

$Y = 1 - X + \epsilon$
$X \sim N(0,1)$
$N = 50$
$M = 3$
$\epsilon \sim N(0,.3)$

APPENDIX A

FORMULAS FOR MISE

A. **X Fixed**

1. $\hat{RQ}1$

$$\text{MISE} = \int_0^1 (\text{BIAS})^2\, dx + \int_0^1 \text{VAR}\,(\hat{R}(x))\, dx$$

$$\int_0^1 (\text{BIAS})^2\, dx = \int_0^1 (E[\hat{RQ}_1] - rQ(u))^2\, du$$

A1. $$= \int_0^1 \left(\frac{\sum_{j=1}^{n} E\left[Y_{[j:n]}\right] K\left(\frac{\frac{j-1}{n-1} - u}{h(n)}\right)}{\sum_{j=1}^{n} K\left(\frac{\frac{j-1}{n-1} - u}{h(n)}\right)} - rQ(u) \right)^2 du$$

where $E\left[Y_{[j:n]}\right] = r\left(X_{(j)}\right)$

$$\int_0^1 \text{VAR}\,(\hat{RQ}_1(u))\, du =$$

$$\int_0^1 \frac{\sum_{j=1}^{n} \text{VAR}\left(Y_{[j:n]}\right) K^2\left(\frac{\frac{j-1}{n-1} - u}{h(n)}\right)}{\left(\sum_{j=1}^{n} K\left(\frac{\frac{j-1}{n-1} - u}{h(n)}\right)\right)^2}\, du$$

Since the ϵ's are uncorrelated

A2. $$= \sigma^2 \int_0^1 \frac{\sum_{j=1}^{n} K^2\left(\frac{\frac{j-1}{n-1} - u}{h(n)}\right)}{\left(\sum_{j=1}^{n} K\left(\frac{\frac{j-1}{n-1} - u}{h(n)}\right)\right)^2}\, du$$

where $\sigma^2 = \text{VAR}(\epsilon)$

2. $\hat{RQ}3$

$$\int_0^1 (\text{Bias})^2\, dx = \int_0^1 \left(\int_{1/2}^{u} E[\hat{RQ}_1'(s)]\, ds + E[\hat{RQ}_1(1/2)] - rQ(u) \right)^2 du$$

where:

A3. $$\int_{1/2}^{u} E[\hat{RQ}_1'(s)]\, ds = \int_{1/2}^{u} \frac{(n-1) \sum_{j=1}^{n-1} E\left(Y_{[j+1:n]}\right) - E\left(Y_{[j:n]}\right) K\left(\frac{\frac{j-1}{n-1} - s}{h(n)}\right)}{\sum_{j=1}^{n-1} K\left(\frac{\frac{j-1}{n-1} - s}{h(n)}\right)}\, ds$$

but when $r(x)$ is a linear function.

$d_j = E\left[Y_{[j+1:n]}\right] - E\left[Y_{[j:n]}\right]$ is a constant d ; therefore A3 becomes

A4. $$(n-1)\, d \int_{1/2}^{u} \frac{\sum_{j=1}^{n-1} K\left(\frac{\frac{j-1}{n-1} - s}{h(n)}\right)}{\sum_{j=1}^{n-1} K\left(\frac{\frac{j-1}{n-1} - s}{h(n)}\right)}\, ds$$ which equals

A5. $$(n-1)\, d\, (u - 1/2)$$

where

A8. $$E\left[Y_{[j:n]}\right] = \int_0^1 rQ(u)\, n\binom{n-1}{j-1} u^{j-1}(1-u)^{n-j}\, du \quad .$$

This was the formula used for calculating the squared bias. As before the trapezoidal rule was used in estimating the integral. Clearly for some distributions of X the integral can not be calculated over the whole range of u. Since 300 intervals were being used in calculating the integral a natural alternative is to calculate the integral over the region $\left(\frac{1}{302}, \frac{301}{302}\right)$. This was used where necessary.

The formula for the variance was computationally difficult and thus the variance was calculated by 100 simulations on each of the models using the trapezoidal rule as before.

$\underline{\hat{RQ}3}$

$$\int_0^1 (\mathrm{BIAS})^2 = \int_0^1 \left(\int_{1/2}^u E[\hat{RQ}_1'(s)]\, ds + E[\hat{RQ}_1(1/2)] - rQ(u) \right)^2 du$$

where

$$\int_{1/2}^u E[\hat{RQ}_1'(s)]\, ds = \int_{1/2}^u (n+1) \frac{\sum_{j=1}^{n-1}\left(E\left[Y_{[j+1:n]}\right] - E\left[Y_{[j:n]}\right]\right) K\left(\frac{\frac{j}{n+1} - s}{h(n)}\right)}{\sum_{j=1}^{n-1} K\left(\frac{\frac{j}{n+1} - s}{h(n)}\right)}\, ds$$

$$= (n+1) \sum_{j=1}^{n-1} \left(E\left[Y_{[j+1:n]}\right] - E\left[Y_{[j:n]}\right]\right) \int_0^1 \frac{K\left(\frac{\frac{j}{n+1} - s}{h(n)}\right)}{\sum_{j=1}^{n-1} K\left(\frac{\frac{j}{n+1} - s}{h(n)}\right)}\, ds$$

where $E\left[Y_{[j:n]}\right]$ is as given in A8 .

Again the variance is computationally difficult so simulations were used here also.

$$E[\hat{RQ}_1(1/2)] = \frac{\sum_{j=1}^{n} E\left[Y_{[j:n]}\right] K\left(\frac{\frac{j-1}{n-1} - \frac{1}{2}}{h(n)}\right)}{\sum_{j=1}^{n} K\left(\frac{\frac{j-1}{n-1} - \frac{1}{2}}{h(n)}\right)}$$

$$\int_0^1 VAR(\hat{RQ}_3)\, du = \int_0^1 \left[VAR\left(\int_{1/2}^{u} \hat{RQ}_1'(s)\, ds\right) + VAR\left(\hat{RQ}_1(1/2)\right) + 2\, COV\left(\int_{1/2}^{u} \hat{RQ}_1'(s)\, ds\, ,\ \hat{RQ}_1(1/2)\right)\right] du$$

where

$$VAR\left(\int_{1/2}^{u} \hat{RQ}_1'(s)\, ds\right) = \int_{1/2}^{u}\int_{1/2}^{u} COV\left(\hat{RQ}_1'(s),\ \hat{RQ}_1'(t)\right) ds\, dt$$

$$= (n-1)^2 \sigma^2 \sum_{j=1}^{n-1} \int_{1/2}^{u} \frac{K\left(\frac{\frac{j-1}{n-1} - s}{h(n)}\right)}{\sum_{j=1}^{n-1} K\left(\frac{\frac{j-1}{n-1} - s}{h(n)}\right)}\, ds$$

$$\times \int_{1/2}^{u} \frac{2K\left(\frac{\frac{j-1}{n-1} - t}{h(n)}\right) - K\left(\frac{\frac{j}{n-1} - t}{h(n)}\right) - K\left(\frac{\frac{j-2}{n-1} - t}{h(n)}\right)}{\sum_{j=1}^{n-1} K\left(\frac{\frac{j-1}{n-1} - t}{h(n)}\right)}\, dt$$

and

$$COV\left(\int_{1/2}^{u} \hat{RQ}_1'(s)\, ds,\ \hat{RQ}_1(1/2)\right) =$$

$$(n-1)\sigma^2 \sum_{j=1}^{n-1} \int_{1/2}^{u} \frac{K\left(\frac{\frac{j-1}{n-1} - s}{h(n)}\right)}{\sum_{j=1}^{n-1} K\left(\frac{\frac{j-1}{n-1} - s}{h(n)}\right)}\, ds \left(\frac{K\left(\frac{\frac{j}{n-1} - \frac{1}{2}}{h(n)}\right) - K\left(\frac{\frac{j-1}{n-1} - \frac{1}{2}}{h(n)}\right)}{\sum_{j=1}^{n} K\left(\frac{\frac{j-1}{n-1} - \frac{1}{2}}{h(n)}\right)}\right)$$

Each of these integrals was evaluated using the trapezoidal rule. The interval (0, 1) was divided into 300 equal segments for purposes of this evaluation.

The integrals in A6 and A7 were done using IMSL subroutine DCADRE with relative error set to be .01 . Since the values of these integrals are small, the error will be very small and an estimate of the error from the subroutine itself was less than .00001 .

X : RANDOM VARIABLE

$\hat{RQ1}$

$$\int_0^1 (BIAS)^2\, dx = \int_0^1 \left(E[\hat{RQ}_1] - rQ(u)\right)^2 du$$

$$= \int_0^1 \left(\frac{\sum_{j=1}^{n} E\left[Y_{[j:n]}\right] K\left(\frac{\frac{j}{n+1} - u}{h(n)}\right)}{\sum_{j=1}^{n} K\left(\frac{\frac{j}{n+1} - u}{h(n)}\right)} - rQ(u)\right)^2 du$$

REFERENCES

Benedetti, J. (1975). "Kernel Estimation of Regression Functions," Proceedings of Computer Science and Statistics: 8th Annual Symposium on the Interface, Health Sciences, Computing Facility, UCLA, pp. 405-408.

Carmichael, J. P. (1978). "Techniques of Quantile Regression," Technical Report No. ARO-5, Statistical Science Division, SUNY/Buffalo.

Parzen, E. (1977). "Nonparametric Statistical Data Science: A Unified Approach Based on Density Estimation and Testing for 'White Noise'," Technical Report No. 47, Statistical Science Division, SUNY/Buffalo.